October 4, 2010

Dr. Ellen Mantus
Senior Program Officer for Risk Analysis
National Academy of Sciences
500 Fifth St., N.W.
Washington, D.C. 20001
Email: emantus@nas.edu

RE: **Summary of Critical Concerns Regarding Reliance on Zhang et al, *Cancer, Epidemiology, Biomarkers & Prevention*, Vol. 19(1), pp 80-88 (2010) as Reported in the Draft IRIS Formaldehyde Assessment**

Dear Members of the NAS Committee:

Prior to 2009, agencies such as IARC (2006) had concluded that the results from epidemiological studies, such as Hauptmann et al. (2003), did not provide sufficient evidence to classify formaldehyde as a known human leukemogen. However, recent assessments by IARC, and by staff of NTP and USEPA, have found that there is now sufficient evidence for a causal association with myeloid leukemia (Baan et al. 2009; NTP 2010a, b; USEPA 2010). While epidemiological studies conducted by Beane Freeman et al. (2009) and Hauptmann et al. (2009) were considered[1], in reaching these findings, much weight, and apparently the deciding factor, has been given to the results provided by Zhang et al. (2010).

As described in the remainder of this letter, there are many serious methodological deficiencies, reporting errors and unsupported biological assumptions in the Zhang et al. (2010) study[2], and these render the conclusions reached by Zhang et al. (2010) at best an over-interpretation of their results, and, in most cases a misinterpretation of their results.

Zhang et al. (2010) purport to measure what they refer to as "leukemia-specific" chromosome changes in CFU-GM cells because CFU cells are the target cells for leukemogenesis and are converted to leukemia stem cells in acute myeloid leukemia (AML)." This is an erroneous assumption based on outdated 30 year old theories of leukemogenesis that are not supported by current scientific and medical knowledge:

A. CFU-GM is not the appropriate assay for measuring clonal events in the development of AML.

- **Cells capable of becoming AML-initiating cells (AML-IC) cannot be measured by CFU-GM.**

[1] Drs. Cole, Mandel, Marsh and Mundt have submitted to this NAS Committee a critique of the primary epidemiological studies. Their review concluded that there is not a causal association with myeloid leukemia because: 1) no statistically significant association between formaldehyde exposure and myeloid was seen for any dose-metric in the Beane Freeman et al. (2009) study, and, 2) methodological limitations in the Hauptmann et al. study call into question the conclusions drawn by the authors.

[2] Individual data for a specific test reported by Zhang et al. (2010) were obtained from NCI in response to a Freedom of Information Act (FOIA) request, and those data provide compelling evidence of these limitations.

The maturation hierarchy of AMLs is analogous to that of cells in normal hematopoiesis. Even in AML, the vast majority of leukemia blasts are endstage cells with limited or no proliferation potential, and they are incapable of self-renewal or propagation of the leukemia. An assay that identifies AML-IC, and that would be useful for demonstrating a clonal lesion in the development of an AML, necessarily must measure a cell capable of self-renewal. In the modern era, the advent of xenograft assays and recombinant cytokine technology has led to a general understanding that AML-IC are less mature than colony-forming cells (e.g. CFU-GM). AML-IC, specifically including those arising in AMLs with aneusomy (e.g. +8, -7), correspond to long-term culture initiating cells (LTC-IC), cobblestone-area forming cells (CAFC), or human severe combined immunodeficient (NOD/SCID) repopulating cells (huSRC)[1-3]. The studies reported by Lapidot et al. (1994), Terpstra et al. (1996) and Ailles et al. (1999) and other studies have demonstrated that AML-IC can be measured using in vitro assays, and that their self-renewal capability is confirmed in vivo. However, AML-IC cannot be measured using CFU-GM or other CFU assays, because such assays do not measure cells with self-renewal capability.

Conditions for culture of huCFU-GM, including those described by Zhang et al (2010), call for incubation for 14 days in methylcellulose with GM-CSF. In contrast, all of the assays for AML-IC require either long term culture (~ 6 weeks) under complex conditions, or transplantation into NOD/SCID mice, in order to identify and measure their proliferative potential. These early cells do not produce differentiated progeny (i.e. colonies) in CFU assays in less than 3 weeks, and they cannot survive under the conditions used for CFU assays, including methylcellulose or agar medium + GM-CSF for CFU-GM[4]. For example, selection and enrichment of AML-IC, using in vitro incubation with agents such as 5-fluorouracil (5-FU), although they do not result in any CFU-GM [2], they do result in enrichment of cells that form CAFC (6 week cultures), and are capable of leukemic transplantation in NOD/SCID mice. Consequently, CAFC and NOD/SCID assays are appropriate for monitoring clonal development of lesions in AML, while CFU-GEMM and CFU-GM are not.

- **Normal human peripheral blood repopulating (stem) cells are vanishingly rare and are not in cell cycle.**

Central to the mechanism of action proposed by Zhang et al. (2010) is the premise that aneusomy occurs in circulating hematopoietic stem cells, which by definition must be dividing, and then return to the bone marrow, the tissue of origin for myeloid leukemias. Although it is impossible to prove a negative, this premise is unsupported. In normal untreated individuals who have not undergone cytokine therapy to mobilize primitive repopulating cells (i.e. stem cells), the frequency of circulating huSRC is so rare as to not be readily measurable [5]. Data on the frequency of circulating stem cells that return to the bone marrow is virtually non-existent, largely because they are so rare that early attempts to use un-mobilized circulating cells to reconstitute hematopoiesis in humans failed ([6-7]). Similarly, the vast majority of circulating repopulating cells in humans are in G_0, and G_0-G_1 progression actually results in depletion of repopulating capacity [8].

B. The methodology used by Zhang et al. (2010) to measure FISH in CFU-GM is seriously inadequate.

- **The reported data do not permit analysis of clonal injury.**

The only advantage of using CFU-GM over liquid culture to measure terminally differentiating myeloid cells is the potential to measure clonal events, i.e. the number of cells containing lesions per colony. However, because Zhang et al. did not report either colony number or the number of colonies scored for fluorescence in situ hybridization (FISH), it is not possible to determine clonal injury.

- **Results of FISH reported by Zhang et al cannot be due to an event occurring in vivo.**

Zhang et al. (2010) report using the CFU-GM protocol published by Stem Cell Technologies, which scores colonies at 14 days containing ≥ 40 cells. This requires a minimum of 6 doublings occurring over a 14-day period to measure a clonal event occurring in a CFU-GM progenitor cell inoculated into culture. A review of the Zhang et al. (2010) primary FISH signal data (n's), obtained from NCI through a Freedom of Information Act (FOIA) request, show that the largest n reported for a FISH signal was 20. The vast majority of n's are ≤ 10 for both the control and "exposed" groups. Based on the kinetics of CFU-GM colony formation [9], all of these signals most likely occurred in or after the 6th division in culture, and almost certainly do not represent events that could have taken place in vivo.

- **The raw data do not comport with the methodology described in Zhang et al. (2010).**

The authors reported that they counted a minimum of 150 cells for each case (exposed = 10, controls = 12). However , the raw data show that far fewer cells were analyzed in the majority of cases. For monosomy 7, a review of the raw data reveals only 1 exposed and 4 control cases for which 150 cells were, in fact, scored. And for the remaining 17 cases, the total number of cells counted ranged from 18-140. (See Appendix 1 to this letter for the table of cells scored for monosomy 7 and trisomy 8.) FISH assays, including those utilizing the specific probes employed by Zhang et al. (2010), are subject to correction for background/sensitivity errors. Because of statistical limitations inherent in the scoring of FISH assays, a minimum of 200 cells (and for certain probes more) are required to report a result in a clinical setting. Statistically significant differences reported for +8 were 1.21% and 0.32% for exposed (n=10) vs unexposed (n=12) subjects. However, if analysis is limited to cases where even ≥ 100 cells are counted, the percentage with +8 is nearly identical (i.e. 1.04% vs 0.94%, respectively).

- **Cutoff values for FISH in normal individuals are not presented.**

Zhang et al (2010) fail to provide cut-off values for +8 and -7 FISH probes in normal individuals. However, in a previous study they indicated an apparent cutoff value of 0.8% ± 0.1% for +8 in controls. [10]. These values call into question whether their FISH analyses can meaningfully resolve the small differences reported in this study.

- **Aneuploidy increases with time in human cells cultured in vitro.**

It is well established that the frequency of aneuploid cells spontaneously increases with time in culture, with significant differences reported in human lymphocytes as early as 72h [11-13]. Further, increases in the frequency of aneuploidy are not random, and numerical changes involving chromosome 8 occur at a higher rate in culture than for many other chromosomes [14].

- **Validation of the CFU-GM/FISH assay to measure aneuploidy in vivo is inadequate**

Previously, Zhang et al. reported that the frequency of aneuploidy in cultured human lymphocytes scored by FISH is several-fold higher in low- and medium-quality metaphase preparations, versus high-quality metaphase preparations [10]. (See Appendix 2 in which Zhang et al. discuss the reliability of their FISH technique specifically for chromosomes 7 and 8.) Given the extended incubation time for CFU-GM used in this study (14d), and the fact that the authors report they scored all available cells, there is a high probability that low- and medium-quality metaphase preparations were counted, leading to an overestimate of aneuploidy.

Notwithstanding that CFU-GM do not measure a repopulating stem cell population, individual variability, inter-testing variability and standard reference ranges should be characterized in a large

normal population prior to the application of this methodology, even for use as a non-specific biomarker of effect.

C. Changes in blood parameters reported by Zhang et al (2010) are not clinically significant.

- **Total counts for WBC, neutrophil, lymphocyte and platelets are well within normal limits for both exposed and unexposed groups.**

- **Statistically significant differences reported between exposed and unexposed groups cannot be attributed to formaldehyde.**

In light of published genetic and regional differences in blood cell counts in Asian populations (see below), insufficient information is provided in Zhang et al. (2010) on the clinical background and origin of exposed individuals (n=43) versus unexposed (n=51) controls to evaluate potential confounding for the small differences between groups reported.

Examples of other potential confounders include:

1) Thalassemia trait.

Although differences are reported for RBC and MCV, Hgb values are identical between exposed and control groups. Analysis of primary data reveals 1 exposed and 4 controls meet criteria (MCV <70fL), suggestive of thalassemia trait. Re-analysis of the data excluding these subjects narrows the difference between MCV and RBC between the two groups, implicating thalassemia trait as a likely confounder.

2) Genetic and regional confounders.

Significant variations in platelet counts in healthy Chinese subjects are known to be influenced by such factors as geographical location, season, and lipid variations [15]. Similarly, genetic polymorphisms have been identified that are associated with significant differences in neutrophil counts in Asian populations [16]. Other influences such as nutrition and Chinese medicine have not been appropriately addressed, and these influences could easily explain the minor variations in CBC parameters that Zhang et al. (2010) instead have attributed to formaldehyde.

- **"Differences" in CFU-GM colonies are not statistically significant.**

Zhang et al describe a "20% decrease in CFU-GM colonies" between exposed [Mean = 7.26%; Range: 1.32-21.38%] and unexposed [Mean = 9.03%; Range:0.84-22.88] groups, which they suggest is due to a toxic effect of formaldehyde even though the results are not statistically significant (p= 0.10). The reported differences in the averages are likely meaningless in view of the fact that there is over a 20-fold variation in the number of colonies in both exposed and unexposed groups.

D. There are no validated mechanisms for the pathogenesis of AML with aneusomy.

- **Neither "Myeloid leukemia" nor AML is a single disease.**

The WHO classification of hematopoietic malignancies stratifies myeloid neoplasms according to broad categories (e.g. chronic myeloproliferative neoplasms, including chronic myelogenous leukemia (CML); myelodysplastic/myeloproliferative diseases (MDS/MPD); myelodysplastic syndromes (MDS); and acute myeloid leukemia (AML)). Insofar as is possible, distinct diseases are defined within each category based on morphology, immunophenotype, genetic abnormalities, clinical features and etiology [17]. Within the category of AML there are several subdivisions comprised of approximately a total of 20 different

disease entities [17-18]. Therefore *a priori* it is not appropriate to consider all myeloid leukemias, or even major subgroups of AML, as a single disease [19].

- **The relevance of monosomy 7 or trisomy 8 as markers for formaldehyde exposure is not established.**

AML, with specific genetic mutations and recurrent cytogenetic abnormalities, are classified on the basis of structural chromosome rearrangements, thereby resulting in chimeric proteins or molecular genetic changes that correlate with alterations in gene expression (gene activation or inactivation). The pathogenesis of these entities, together with their prognostic significance, are well defined and form the basis for classification of CML and AML with reoccurring cytogenetic abnormalities [17].

Zhang et al. (2010) refer to their identification of monosomy 7 and trisomy 8 as "Leukemia-Specific" chromosome changes. However, there is no mention in Zhang et al. (2010) regarding analysis for potential aneuploidy of other chromosomes. Because there are no chromosome studies available for the leukemia cases reported in either controls or in exposed individuals in the available epidemiologic studies, it is not possible to ascertain whether monosomy 7 or trisomy 8 are relevant markers to study in formaldehyde exposed workers.

- **Benzene-induced MDS/AML is not a positive control for aneusomy occurring in the pathogenesis of hematopoietic neoplasms.**

Over the past 20 years, hypotheses for a role for aneuploidy in the development of benzene-induced hematopoietic neoplasms have been the subject of numerous studies, including several authored by one of us. However, these hypotheses are not supported by the results of recent epidemiology studies employing state-of-the-art molecular and cytogenetic methods. Although current studies confirm previous observations that benzene plays a causal role in development of subtypes of MDS and AML, those studies do not provide any evidence for aneusomy in the development of these diseases [20-21] .

In conclusion, based on the fundamental biological misconceptions, methodological deficiencies and inaccuracies described above, we believe that the Zhang et al. (2010) study cannot be considered reliable, and it should not be used as a basis to confirm or suggest a relationship between formaldehyde and AML or any other leukemia.

Sincerely,

Richard D. Irons, MT, PhD, DABT
Professor Emeritus of Pathology and Toxicology
University of Colorado at Denver
Cinpathogen International, LLC
Boulder, CO

Michael J. Thirman, MD
Associate Professor of Medicine
Director of Leukemia Biology
Hematology/Oncology
University of Chicago

Richard J. Albertini, MD, PhD
Professor Emeritus, Medicine
Research Professor, Pathology
University of Vermont College of Medicine
Burlington, VT

Annette M. Shipp, PhD
Principal
ENVIRON International Corporation
Monroe, LA

References

1. Lapidot, T., et al., *A cell initiating human acute myeloid leukaemia after transplantation into SCID mice.* Nature 1994. **367**: p. 645-648.
2. Terpstra, W., et al., *Fluorouracil selectively spares acute myeloid leukemia cells with long-term growth abilities in immunodeficient mice and in culture.* Blood, 1996. **88**: p. 1944-1950.
3. Ailles, L., et al., *Growth characteristics of acute myelogenous leukemia progenitors that initiate malignant hematopoiesis in nonobese diabetic/severe combined immunodeficient mice.* Blood, 1999. **94**(5): p. 1761-1772.
4. Coulombel, L., *Identification of hematopoietic stem/progenitor cells: strength and drawbacks of functional assays.* Oncogene, 2004. **23**: p. 7210-7222.
5. Wang, J., M. Doedens, and J. Dick, *Primitive human hematopoietic cells are enriched in cord blood compared with adult bone marrow or mobilized peripheral blood as measured by the quantitative in vivo SCID-repopulating cell assay.* Blood, 1997. **89**(11): p. 3919-3924.
6. Abrams, R., et al., *Result of attempted hematopoietic reconstitution using isologous, peripheral blood mononuclear cells: a case report.* Blood, 1980. **56**(3): p. 516-520.
7. Hershko, C., et al., *Cure of aplastic anemia in paroxysmal nocturnal haemoglobinuria by marrow transfusion from identical twin: failure of peripheral blood-leukocyte transfusion to correct marrow aplasia.* Lancet, 1979. **1**: p. 945-947.
8. Gothot, A., et al., *Cell cycle-related changes in repopulating capacity of human mobilized peripheral blood CD34+ cells in non-obese diabetic/severe combined immune-deficient mice.* Blood, 1998. **92**(8): p. 2641-2649.
9. Lewis, J., N. Blackett, and M. Gordon, *The kinetics of colony formation by CFU-GM in vitro.* British J of Haematology, 1994. **88**: p. 440-442.
10. Zhang, L., et al., *Benzene increases aneuploidy in the lymphocytes of exposed workers: a comparison of data obtained by fluorescence in situ hybridization in interphase and metaphase cells.* Environmental & Molecular Mutagenesis, 1999. **34**(4): p. 260-8.
11. Richard, F., et al., *Aneuploidy in human lymphocytes: an extensive study of eight individuals of various ages.* Mutation Research, 1993. **295**: p. 71-80.
12. Sherwood, S., et al., *Defining cellular senescence in IMR-90 cells: A flow cytometric analysis.* Proc Natl Acad Sci, USA, 1988. **85**: p. 9086-9090.
13. Zotova, N., et al., *Spontaneous aneuploidy level in blood cells of fertile females.* Cell and Tissue Biology, 2009. **51**(7): p. 585-591.
14. Mukherjee, A.B. and S. Thomas, *A longitudinal study of human age-related chromosomeal analysis in skin fibroblasts.* Exptl Cell Research, 1997. **235**: p. 161-169.
15. Peng, L., et al., *Effects of biological variations on platelet count in healthy subjects in China.* Thromb. Haemost., 2004. **91**: p. 367-372.
16. Okada, Y., et al., *Common variations in PSMD3-CSF3 and PLCB4 are associated with neutrophil count.* Human Molecular Genetics, 2010. **19**(10): p. 2079-2085.
17. Vardiman, J., et al., *Introduction and overview of the classification of myeloid neoplasms.*, in *WHO Classification of Tumours of the Haematoipoietic and Lymphoid Tissues*, S. Swerdlow, et al., Editors. 2008, International Agency for Research on Cancer (IARC): Lyon. p. 18-30.
18. Harris, N., et al., *WHO classification of tumours of haematopoietic and lymphoid tissues: Introduction*, in *World Health Organization Classification of Tumours*, E. Jaffee, et al, Editor. 2001, IARC Press: Lyon, France. p. 12-13.

19. Appelbaum, F., et al., *Acute myeloid leukemia.* Hematology, 2001. **2001**: p. 62-86.
20. Irons, R.D., et al., *Integrating WHO 2001-2008 criteria for the diagnosis of Myelodysplastic Syndrome (MDS): a case-case analysis of benzene exposure.* Chem Biol Interact, 2010. **184**: p. 30-38.
21. Wong, O., et al., *A hospital-based case-control study of acute myeloid leukemia in Shanghai: analysis of environmental and occupational risk factors by subtypes of the WHO classification.* Chemico-Biological Interactions, 2010. **184**: p. 112-128.

Appendix 1
Levels of Monosomy of Chromosome 7 and Trisomy of Chromosome 8 in cells Scored by (Zhang et al. 2010)

Monosomy 7				Trisomy 8			
Exposed		Unexposed		Exposed		Unexposed	
Number Detected	Number Analyzed	Number Detected	Number Analyzed	Number Detected	Number Analyzed	Number Detected	Number Analyzed
11	274	19	288	2	192	2	226
15	132	10	272	4	180	2	215
20	123	10	260	4	173	2	197
4	109	8	163	1	149	1	94
4	101	6	140	0	139	0	91
3	95	2	78	0	108	0	83
9	76	1	71	2	78	0	69
13	61	9	70	2	61	0	67
10	50	4	49	0	53	0	37
6	39	0	24	0	33	0	25
		2	20			0	22
		1	18			0	21

Appendix 2
Reproduced from Zhang et al. (1999), p. 266 [10]

"One issue that must be addressed is the apparently high rates of monosomy (cells with one hybridization signal) and trisomy (cells with three hybridization signals) reported in metaphase cells in the present study compared with rates obtained by classical cytogenetics. The rates of apparent monosomy and trisomy most likely result from the fact that we examined all scorable metaphases on the slides, as previously defined [Zhang et al, 1998b], rather than just the 50 best. This approach accounts for the differences between our data and the much lower numbers generated by conventional analysis of a limited number of high quality metaphase spreads. When we reexamined around 50 of the best metaphase spreads by FISH in three subjects with high aneuploidy rates, we detected very few aneuploid cells. In fact, on average, the rate of aneuploidy was several-fold lower in the best 50 metaphase spreads compared with those of lower quality (data not shown). Therefore, using FISH to analyze only the best quality spreads would have yielded a lower aneusomy rate comparable to rates obtained by classical cytogenetics. The high values we report here and elsewhere result from our scoring poor- and medium- as well as high quality spreads. This approach appears to increase our power to detect chromosomal damage in exposed populations, but does not allow for ready comparison of our data with measurements of aneuploidy by conventional cytogenetics or even by FISH in the best metaphase spreads."